Quantifying Intramolecular Binding in Multivalent Interactions: A Structure-based Synergistic Study on the Grb2-Sos1 Complex

Anurag Sethi, Byron Goldstein, S. Gnanakaran, T-6

The formation of multivalent complexes by combining relatively weak promiscuous interactions is a strategy that is often used to increase the affinity and specificity of biomolecular complex formation among signaling proteins. However, there is often no clear structural information regarding multivalent interactions between signaling protein partners and their effect on the stoichiometry and the distribution of these protein complexes. We developed a hybrid method that combines molecular dynamics simulations and polymer models to estimate the intramolecular concentration of a motif near the vacant binding site in the same complex during multivalent binding. We developed and applied this method to the study of complex formation between the adaptor protein growth receptor binding protein-2 (Grb2) and son of sevenless-1 (Sos1). Using this method, we were able to estimate the stoichiometry and distribution of complexes formed under physiological conditions.

Many signaling proteins use multivalent interactions in which relatively weak interactions are combined to increase the affinity and specificity of complex formation. Often, in such interactions, one of the biomolecules (protein A) involved consists of a disordered region that contains multiple ligands (or motifs) that can each bind to the binding sites of a protein partner (protein B). Often the protein, in our case growth receptor bound protein-2 (Grb2), is bivalent. When one of the binding sites of protein B is bound to a motif on protein B, the two molecules are tethered to each other. This results in an increase of the local concentration of the free motifs on protein B near the free site(s) on protein A. The related intramolecular binding constants and effective concentrations are very difficult to determine experimentally. In order to estimate the intramolecular binding constant(s) theoretically, polymer models are often used to model the disordered regions in the biomolecule. However, the flexibility of the protein that contains the binding site(s) is ignored. We developed a hybrid approach combining molecular dynamics simulations and polymer modeling that accounts for the flexibility in both molecules to estimate the intramolecular binding constants [1].

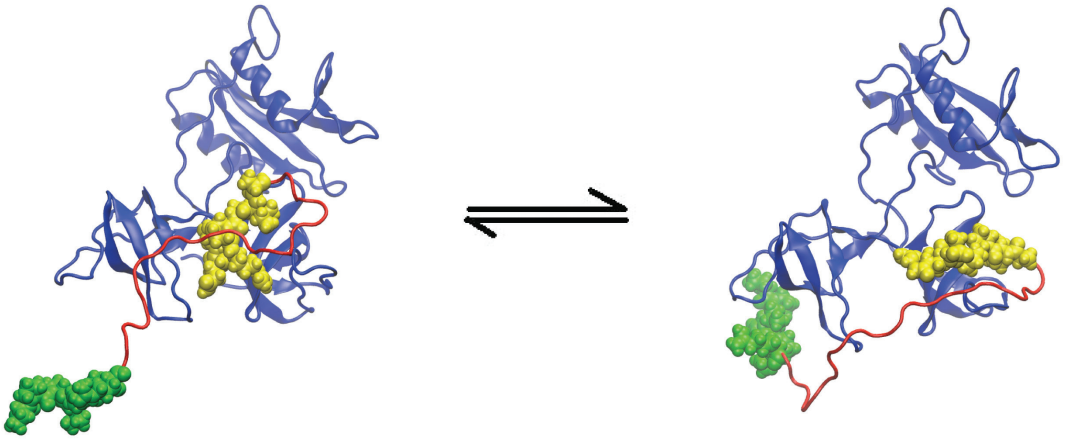

Fig. 1. The intramolecular binding of a polyproline motif (green) on Sos1 to the SH3 domain of Grb2 (blue) after another polyproline motif (yellow) in the same Sos1 molecule is bound. The two polyproline motifs in Sos1 are connected by a linker (red).

We applied this method to the study of the multivalent binding between the adaptor protein Grb2 and son of sevenless-1 (Sos1) and provided a rationale for the stoichiometry of the complexes that were observed experimentally. Grb2 contains two SH3 domains (small protein domain of about 60 amino acid residues) that interact with multiple polyproline motifs separated by flexible unstructured regions on Sos1. Grb2 mediates the recruitment of Sos1 from the cytosol to the plasma membrane where it activates Ras (a family of related proteins involved in transmitting signals within cells) by inducing the exchange of guanosine diphosphate (GDP) for guanosine triphosphate (GTP). First, using a combination of evolutionary information and binding-energy calculations, we predict an additional polyproline motif in Sos1 that binds to the SH3 domains of Grb2. This gives rise to a total of five polyproline motifs in Sos1 that are capable of binding to the two SH3 domains of Grb2. Then, combining molecular dynamics simulations and polymer models, we estimate the enhancement in local concentration of a polyproline motif on Sos1 near an unbound SH3 domain of Grb2 when

For more information contact S. Gnanakaran at gnana@lanl.gov

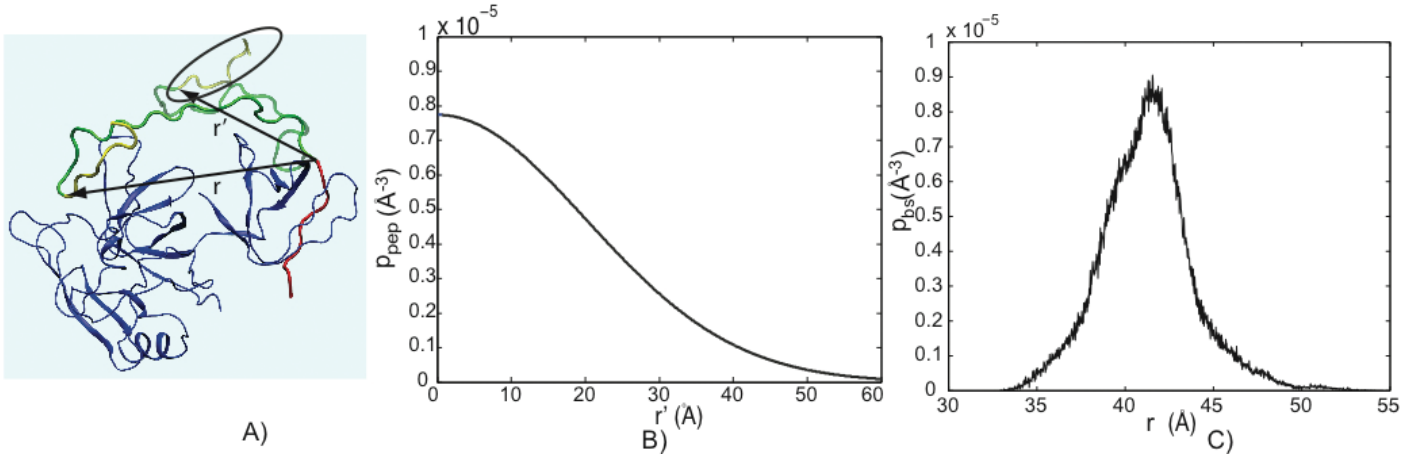

Fig. 2. (a) An SH3 domain on Grb2 (blue) is exposed to a greater concentration of a polyproline motif (yellow) on Sos1 (circled) after one of the motifs (red) in Sos1 is bound to the other SH3 domain in Grb2. The linker region connecting the two motifs is shown in green. In the hybrid model, we use a (b) worm-like chain model to calculate the probability density of the distance between the two ends of the linker, while a (c) molecular dynamics simulation is used to measure the probability density of the distance between the two binding sites in Grb2.

its other SH3 domain is bound to a different polyproline motif on Sos1. We show that the local concentration of the Sos1 motifs that a Grb2 SH3 domain experiences is approximately 1000 times greater than the cellular concentration of Sos1. Finally, we calculate the intramolecular equilibrium constants for the crosslinking of Grb2 on Sos1 and use thermodynamic modeling to calculate the stoichiometry. With these equilibrium constants, we were able to predict the distribution of complexes that form at physiological concentrations. We believe this is the first systematic analysis that combines sequence, structure, and thermodynamic analyses to determine the stoichiometry of the complexes that are dominant in the cellular environment.

This is a general method that can be used to model many multivalent complexes observed in biology, such as in signaling molecules, antibody-antigen binding, host-pathogen interactions, and DNA-transcription factor interactions. In addition to estimating the intramolecular binding constants, we also gained a molecular understanding of the nature of complexes that are formed in these systems. These principles could be used for rational design of drugs or vaccines that would optimally bind in a multivalent fashion based upon such models.

[1] Sethi, A. et al., *PLoS Comput Biol* **7**, e1002192 (2011).

Funding Acknowledgments
National Institutes of Health; LANL Institutional Computing Program; LANL Laboratory Directed Research and Development Program